Sparse Signal Reconstruction with a Sign Oracle

Arthur Marmin*, Marc Castella†, Jean-Christophe Pesquet*
* *Center for Visual Computing, CentraleSupélec, INRIA, Université Paris-Saclay,* Gif-sur-Yvette, France
{arthur.marmin, jean-christophe.pesquet}@centralesupelec.fr
† *SAMOVAR, Télécom SudParis, CNRS Université Paris-Saclay,* Evry, France
marc.castella@telecom-sudparis.eu

***Abstract*—Sparse signal reconstruction is performed by minimizing the sum of a least-squares fit regularized with a piecewise rational approximation of ℓ_0. We show the benefit of an oracle that yields the sign of the signal when using a recent methodology for global polynomial or semi-algebraic minimization. The computational time and memory cost are both decreased.**

I. Introduction

With the exponential growth of numerical data acquisition and storage requirements, compressive sensing has become a fundamental tool. The standard approach to promote sparse solutions consists in adding an ℓ_0 penalization to a data-fit cost function [1]. Since this yields NP hard optimization problems, surrogates for ℓ_0 that are continuous, exact and that do not add undesirable bias were proposed [2]–[7]. Nevertheless, they are non convex and lead to intricate optimization problems which can be solved only approximately [5], [8]–[13]. Fortunately, many criteria are piecewise rational and the problem can then be reformulated in terms of polynomial optimization problems before being solved exactly following a recent approach [14]–[16]. In this work, we explore the advantage provided by an a priori knowledge on the sign of the estimated signal.

II. Problem modelling and methodology

We consider the reconstruction of an unknown sparse discrete-time signal $\overline{\mathbf{x}} \in \mathbb{R}^T$ from observations $\mathbf{y} \in \mathbb{R}^T$ given by the following linear degradation model $\mathbf{y} = \mathbf{H}\overline{\mathbf{x}} + \mathbf{w}$, where $\mathbf{w} \in \mathbb{R}^T$ is a zero-mean white Gaussian noise and $\mathbf{H} \in \mathbb{R}^{T\times T}$ is a Toeplitz band matrix associated to a convolution filter of length L. We estimate $\overline{\mathbf{x}}$ as a minimizer of the following regularized criterion

$$(\forall \mathbf{x} \in \mathbb{R}^T) \quad \mathcal{J}(\mathbf{x}) = \frac{1}{2}\left\|\mathbf{y} - \mathbf{H}\mathbf{x}\right\|^2 + \sum_{t=1}^{T} \Psi_\lambda(|x_t|)\,, \qquad (1)$$

where Ψ_λ is a piecewise rational regularization that promotes sparsity and depends on a parameter $\lambda \in]0, +\infty[$. Many continuous exact relaxations of ℓ_0 can be expressed in the above form [2]–[7]. In particular, the penalization is generally an even separable function which depends on $|x_t|$ only and can be written $\Psi_\lambda(|x_t|)$. The minimization of (1) is first reformulated as a polynomial optimization problem and then solved using Lasserre's method [16]. Our methodology is applicable to any piecewise rational function. Since $\overline{\mathbf{x}}$ is a real-valued signal, in our previous work, we used additional variables $\mathbf{u} \in \mathbb{R}^T$ such that $\mathbf{u}^2 = |\mathbf{x}|^2$, $\mathbf{u} \geq 0$ as a substitute for the absolute value in Ψ_λ.

Lasserre's method constructs a hierarchy of convex SDP problems indexed by the relaxation order k. Solving each SDP problem yields both a lower bound $\mathcal{J}_k^*$ on the criterion $\mathcal{J}$ and a minimizer $\hat{\mathbf{x}}$ [17]. $(\mathcal{J}_k^*)_{k\in\mathbb{N}}$ is an increasing convergent sequence whose limit is the true optimal value of (1) [14].

III. Use of an oracle on the sign of $\overline{\mathbf{x}}$

For real-valued signals, convergence is observed for orders k for which building and solving the corresponding SDP problems is highly demanding in terms of computation and memory storage. However, when $\overline{\mathbf{x}}$ is a positive signal, we observe [16] convergence at a lower order k. This thus suggests a method yielding similar results for real-valued signals using an oracle. Instead of (1), we minimize

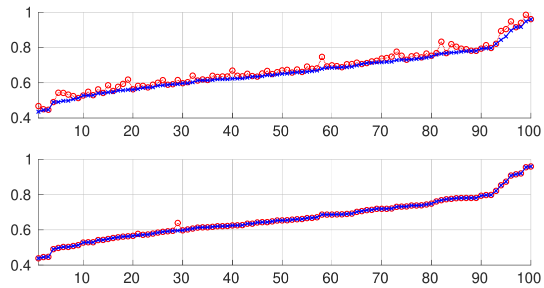
$$(\forall \mathbf{v} \in \mathbb{R}_+^T) \quad \widetilde{\mathcal{J}}(\mathbf{v}) = \frac{1}{2}\left\|\mathbf{y} - \widetilde{\mathbf{H}}\mathbf{v}\right\|^2 + \sum_{t=1}^{T} \Psi_\lambda(v_t)\,,$$

where $\mathbf{v} \in \mathbb{R}_+^T$, $\widetilde{\mathbf{H}} = \mathbf{H}\,\mathrm{Diag}(\boldsymbol{\epsilon})$, $\boldsymbol{\epsilon} \in \{-1, 1\}^T$ is the sign vector of $\overline{\mathbf{x}}$ provided by the oracle, and $\mathrm{Diag}(\boldsymbol{\epsilon})$ is the scalar matrix with diagonal elements $\boldsymbol{\epsilon}$. We build our oracle by solving (1) with lasso [18], i.e. $\Psi_\lambda = \lambda\,\mathrm{Id}$. The availability of an oracle allows us to restrict the minimization of (1) to positive signals thanks to the new convolution matrix $\widetilde{\mathbf{H}}$. An exact solution is thus retrieved by solving a SDP problem of fair dimension. Moreover, the use of an oracle decreases significantly the computation time since additional variables $\mathbf{u}$ are no longer required.

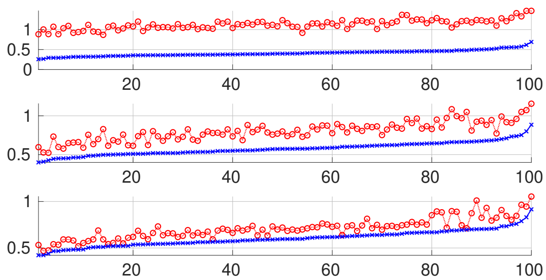
IV. Numerical simulations

We generate 100 Monte-Carlo realizations of the signal $\overline{\mathbf{x}}$ with size $T = 50$ and 100 and 10% non-zero elements drawn randomly from the uniform distribution on $[-1, -2/3] \cup [2/3, 1]$. The length L of the filter is set to 3. We use the penalization SCAD [5]. We solve the SDP corresponding to relaxation orders 2, 3, and 4.

Figure 1 shows the value of the criterion at the minimizer and the lower bound for different cases. When the oracle is available, we observe that the convergence is reached at $k = 3$ since $\widetilde{\mathcal{J}}_k^* = \widetilde{\mathcal{J}}(\hat{\mathbf{x}}_{\text{our oracle}})$. Conversely, when no information on the sign of $\overline{\mathbf{x}}$ is available, there is still a gap between $\mathcal{J}_k^*$ and $\mathcal{J}(\hat{\mathbf{x}}_{\text{no oracle}})$ at $k = 4$. Figure 2 shows the quality of our oracle relative to the original criterion (1). Table I shows the corresponding average computational times. Using our oracle secures convergence for a lower relaxation order and thus leads to a significant decrease of computational time. Table II shows the relative error for the different estimated signals $\hat{\mathbf{x}}$.



(a) With our oracle: in red $\widetilde{\mathcal{J}}(\hat{\mathbf{x}}_{\text{our oracle}})$, in blue $\widetilde{\mathcal{J}}_k^*$ for $k = 2$ (top), 3 (bottom)



(b) Without oracle: in red $\mathcal{J}(\hat{\mathbf{x}}_{\text{no oracle}})$, in blue $\mathcal{J}_k^*$ for $k = 2$ (top), 3 (middle), and 4 (bottom)

Fig. 1. Comparison between the lower bound and the optimal value of the criterion for 100 tests with $T = 50$. The superposition of the two curves indicates the convergence of Lasserre's hierarchy. For the sake of clarity, the cases are ordered by increasing value of the lower bound.

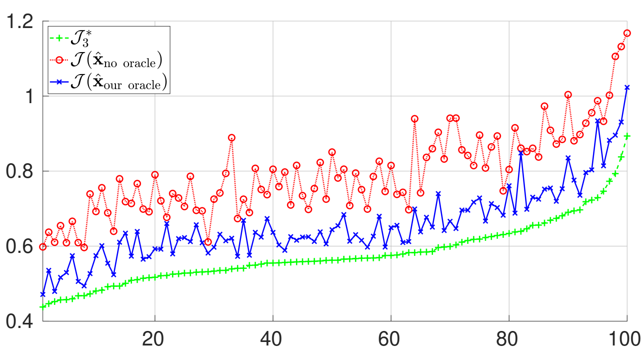


Fig. 2. Value of the original criterion $\mathcal{J}(\hat{\mathbf{x}})$ for the different estimates $\hat{\mathbf{x}}$ on 100 tests with $T = 50$ and $k = 3$. For the sake of clarity, the cases are ordered by increasing value of $\mathcal{J}_3^*$.

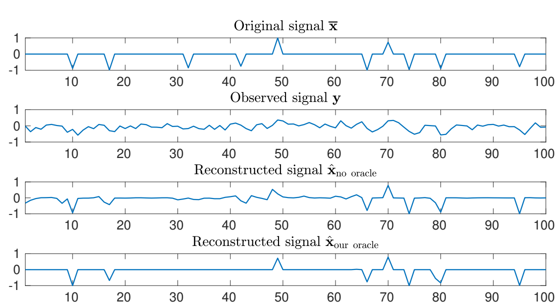


Fig. 3. Comparison of the different estimates of $\overline{\mathbf{x}}$ ($T = 100$ and $k = 3$)

TABLE I
AVERAGE COMPUTATIONAL TIMES

	$T = 50$		$T = 100$	
	Our oracle	No oracle	Our oracle	No oracle
$k = 2$	9s	9s	15s	16s
$k = 3$	19s	25s	46s	1min
$k = 4$	15min53s	27min38s	2h11s	3h18min43s

TABLE II
AVERAGE RELATIVE RECONSTRUCTION ERROR

$\frac{\|\hat{\mathbf{x}}-\overline{\mathbf{x}}\|}{\|\overline{\mathbf{x}}\|}$	No oracle	Our oracle	Perfect oracle	Lasso
$k = 2$	0.765	0.694	0.545	0.809
$k = 3$	0.710	0.710	0.564	

REFERENCES

[1] M. Nikolova, "Description of the minimizers of least squares regularized with ℓ_0 norm. Uniqueness of the global minimizer," *SIAM J. Imaging Sci.*, vol. 6, no. 2, pp. 904–937, Jan. 2013.

[2] T. Zhang, "Analysis of multi-stage convex relaxation for sparse regularization," *J. Mach. Learn. Res.*, vol. 11, pp. 1081–1107, Mar. 2010.

[3] M. Artina, M. Fornasier, and F. Solombrino, "Linearly constrained nonsmooth and nonconvex minimization," *SIAM J. Optim.*, vol. 23, no. 3, pp. 1904–1937, Jan. 2013.

[4] A. Jezierska, H. Talbot, O. Veksler, and D. Wesierski, "A fast solver for truncated-convex priors: Quantized-convex split moves," in *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 45–58.

[5] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *J. Am. Stat. Assoc.*, vol. 96, no. 456, pp. 1348–1360, Dec. 2001.

[6] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Ann. Appl. Stat.*, vol. 38, no. 2, pp. 894–942, Apr. 2010.

[7] E. Soubies, L. Blanc-Féraud, and G. Aubert, "A continuous exact ℓ_0 penalty (CEL0) for least squares regularized problem," *SIAM J. Imaging Sci.*, vol. 8, no. 3, pp. 1607–1639, jan 2015.

[8] P. Ochs, A. Dosovitskiy, T. Brox, and T. Pock, "On iteratively reweighted algorithms for nonsmooth nonconvex optimization in computer vision," *SIAM J. Imaging Sci.*, vol. 8, no. 1, pp. 331–372, Jan. 2015.

[9] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 877–905, Oct. 2008.

[10] P. Breheny and J. Huang, "Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection," *Ann. Appl. Stat.*, vol. 5, no. 1, pp. 232–253, Mar. 2011.

[11] A. Patrascu and I. Necoara, "Random coordinate descent methods for ℓ_0 regularized convex optimization," *IEEE Trans. Automat. Contr.*, vol. 60, no. 7, pp. 1811–1824, July 2015.

[12] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 629–654, Sept. 2008.

[13] I. Selesnick, "Sparse regularization via convex analysis," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4481–4494, Sept. 2017.

[14] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM J. Optim.*, vol. 11, no. 3, pp. 796–817, Jan. 2001.

[15] F. Bugarin, D. Henrion, and J. B. Lasserre, "Minimizing the sum of many rational functions," *Math. Program. Comput.*, vol. 8, no. 1, pp. 83–111, Aug. 2015.

[16] M. Castella, J.-C. Pesquet, and A. Marmin, "Rational optimization for nonlinear reconstruction with approximate ℓ_0 penalization," *IEEE Trans. Signal Process.*, vol. 67, pp. 1–1, 2018.

[17] D. Henrion and J.-B. Lasserre, "Detecting global optimality and extracting solutions in GloptiPoly," in *Positive Polynomials in Control*. Springer Berlin Heidelberg, Sept. 2005, vol. 312, pp. 293–310.

[18] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 58, no. 1, pp. 267–288, 1996.